Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	5	(("6,834,109") or ("6,647,341") or ("6, 084,547") or ("5,917,449") or ("5,717, 406")).PN.	US-PGPUB; USPAT	OR	OFF	2005/02/05 19:49

IEEE HOME | SEARCH IEEE | SHOP | WEB ACCOUNT | CONTACT IEEE

IEEE

Membership Publications/Services Standards Conferences Careers/Jobs

IEEE Xplore® RELEASE 1.8

Welcome
United States Patent and Trademark Office

IEEE Xplore®
1 Million Documents
1 Million Users
...And Growing

» Search Results

Help FAQ Terms IEEE Peer Review Quick Links

Welcome to IEEE Xplore®
- Home
- What Can I Access?
- Log-out

Tables of Contents
- Journals & Magazines
- Conference Proceedings
- Standards

Search
- By Author
- Basic
- Advanced
- CrossRef

Member Services
- Join IEEE
- Establish IEEE Web Account
- Access the IEEE Member Digital Library

IEEE Enterprise
- Access the IEEE Enterprise File Cabinet

Your search matched **1** of **1123491** documents.
A maximum of **500** results are displayed, **15** to a page, sorted by **Relevance** in **Descending** order.

Refine This Search:
You may refine your search by editing the current search expression or entering a new one in the text box.

k means<and>predict<and>margin Search

☐ Check to search within this result set

Results Key:
JNL = Journal or Magazine **CNF** = Conference **STD** = Standard

1 **On data based learning using support vector clustering**
Ribeiro, B.;
Neural Information Processing, 2002. ICONIP '02. Proceedings of the 9th International Conference on , Volume: 5 , 18-22 Nov. 2002
Pages:2516 - 2521 vol.5

[Abstract] [PDF Full-Text (636 KB)] IEEE CNF

Print Format

Home | Log-out | Journals | Conference Proceedings | Standards | Search by Author | Basic Search | Advanced Search | Join IEEE | Web Account | New this week | OPAC Linking Information | Your Feedback | Technical Support | Email Alerting | No Robots Please | Release Notes | IEEE Online Publications | Help | FAQ | Terms | Back to Top

Copyright © 2004 IEEE — All rights reserved

ACM PORTAL
US Patent & Trademark Office
Subscribe (Full Service) Register (Limited Service, Free) Login
Search: The ACM Digital Library The Guide
"k-means" and predict and margin
SEARCH
Feedback Report a problem Satisfaction survey

Terms used **k means** and **predict** and **margin** Found **3,433** of **150,138**

Sort results by: relevance
Display results: expanded form
Save results to a Binder
Search Tips
Open results in a new window

Try an Advanced Search
Try this search in The ACM Guide

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next
Best 200 shown Relevance scale

1 Solving regression problems with rule-based ensemble classifiers
Nitin Indurkhya, Sholom M. Weiss
August 2001 **Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining**
Full text available: pdf(556.71 KB) Additional Information: full citation, abstract, references, index terms

We describe a lightweight learning method that induces an ensemble of decision-rule solutions for regression problems. Instead of direct prediction of a continuous output variable, the method discretizes the variable by k-means clustering and solves the resultant classification problem. Predictions on new examples are made by averaging the mean values of classes with votes that are close in number to the most likely class. We provide experimental evidence that this indirect approach can often yi ...

2 Boosting margin based distance functions for clustering
Tomer Hertz, Aharon Bar-Hillel, Daphna Weinshall
July 2004 **Twenty-first international conference on Machine learning**
Full text available: pdf(181.49 KB) Additional Information: full citation, abstract, references

The performance of graph based clustering methods critically depends on the quality of the distance function used to compute similarities between pairs of neighboring nodes. In this paper we learn distance functions by training binary classifiers with margins. The classifiers are defined over the product space of pairs of points and are trained to distinguish whether two points come from the same class or not. The signed margin is used as the distance value. Our main contribution is a distance l ...

3 Online and batch learning of pseudo-metrics
Shai Shalev-Shwartz, Yoram Singer, Andrew Y. Ng
July 2004 **Twenty-first international conference on Machine learning**
Full text available: pdf(192.62 KB) Additional Information: full citation, abstract, references

We describe and analyze an online algorithm for supervised learning of pseudo-metrics. The algorithm receives pairs of instances and predicts their similarity according to a pseudo-metric. The pseudo-metrics we use are quadratic forms parameterized by positive semi-definite matrices. The core of the algorithm is an update rule that is based on successive projections onto the positive semi-definite cone and onto half-space constraints imposed by the examples. We describe an efficient procedure fo ...

4 Feature Selection for Unsupervised Learning

Jennifer G. Dy, Carla E. Brodley
August 2004 **The Journal of Machine Learning Research**, Volume 5

Full text available: pdf(725.21 KB) Additional Information: full citation, abstract

In this paper, we identify two issues involved in developing an automated feature subset selection algorithm for unlabeled data: the need for finding the number of clusters in conjunction with feature selection, and the need for normalizing the bias of feature selection criteria with respect to dimension. We explore the feature selection problem and these issues through FSSEM (Feature Subset Selection using Expectation-Maximization (EM) clustering) and through two different performance criteria ...

5 Extracting predicates from mining models for efficient query evaluation
Surajit Chaudhuri, Vivek Narasayya, Sunita Sarawagi
September 2004 **ACM Transactions on Database Systems (TODS)**, Volume 29 Issue 3

Full text available: pdf(698.37 KB) Additional Information: full citation, abstract, references, index terms

Modern relational database systems are beginning to support ad hoc queries on mining models. In this article, we explore novel techniques for optimizing queries that contain predicates on the results of application of mining models to relational data. For such queries, we use the internal structure of the mining model to automatically derive traditional database predicates. We present algorithms for deriving such predicates for a large class of popular discrete mining models: decision trees, nai ...

Keywords: Complex predicate optimization, simpler rules from complex predictive functions

6 Clustering: Restrictive clustering and metaclustering for self-organizing document collections
Stefan Siersdorfer, Sergej Sizov
July 2004 **Proceedings of the 27th annual international conference on Research and development in information retrieval**

Full text available: pdf(171.71 KB) Additional Information: full citation, abstract, references, index terms

This paper addresses the problem of automatically structuring heterogenous document collections by using clustering methods. In contrast to traditional clustering, we study restrictive methods and ensemble-based meta methods that may decide to leave out some documents rather than assigning them to inappropriate clusters with low confidence. These techniques result in higher cluster purity, better overall accuracy, and make unsupervised self-organization more robust. Our comprehensive experimenta ...

Keywords: meta clustering, restrictive clustering

7 Unsupervised Bayesian visualization of high-dimensional data
Petri Kontkanen, Jussi Lahtinen, Petri Myllymäki, Henry Tirri
August 2000 **Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available: pdf(160.91 KB) Additional Information: full citation, references, index terms

8 Context-specific Bayesian clustering for gene expression data
Yoseph Barash, Nir Friedman
April 2001 **Proceedings of the fifth annual international conference on Computational biology**

Full text available: pdf(233.32 KB) Additional Information: full citation, abstract, references, citings, index terms

The recent growth in genomic data and measurement of genome-wide expression patterns allows to examine gene regulation by transcription factors using computational tools. In this work, we present a class of mathematical models that help in understanding the connections between transcription factors and functional classes of genes based on genetic and genomic data. These models represent the joint distribution of transcription factor binding sites and of expression levels of a gene in a single ...

9 Data mining: A matrix density based algorithm to hierarchically co-cluster documents and words

Bhushan Mandhani, Sachindra Joshi, Krishna Kummamuru
May 2003 **Proceedings of the twelfth international conference on World Wide Web**

Full text available: pdf(133.06 KB) Additional Information: full citation, abstract, references, citings, index terms

This paper proposes an algorithm to hierarchically cluster documents. Each cluster is actually a cluster of documents and an associated cluster of words, thus a document-word co-cluster. Note that, the vector model for documents creates the document-word matrix, of which every co-cluster is a submatrix. One would intuitively expect a submatrix made up of high values to be a good document cluster, with the corresponding word cluster containing its most distinctive features. Our algorithm looks to ...

10 A hierarchical method for multi-class support vector machines

Volkan Vural, Jennifer G. Dy
July 2004 **Twenty-first international conference on Machine learning**

Full text available: pdf(171.20 KB) Additional Information: full citation, abstract, references

We introduce a framework, which we call Divide-by-2 (DB2), for extending support vector machines (SVM) to multi-class problems. DB2 offers an alternative to the standard one-against-one and one-against-rest algorithms. For an N class problem, DB2 produces an $N - 1$ node binary decision tree where nodes represent decision boundaries formed by $N - 1$ SVM binary classifiers. This tree structure allows us to present a generalization and a time complexity analysis of DB ...

11 Data clustering: a review

A. K. Jain, M. N. Murty, P. J. Flynn
September 1999 **ACM Computing Surveys (CSUR)**, Volume 31 Issue 3

Full text available: pdf(636.24 KB) Additional Information: full citation, abstract, references, citings, index terms, review

Clustering is the unsupervised classification of patterns (observations, data items, or feature vectors) into groups (clusters). The clustering problem has been addressed in many contexts and by researchers in many disciplines; this reflects its broad appeal and usefulness as one of the steps in exploratory data analysis. However, clustering is a difficult problem combinatorially, and differences in assumptions and contexts in different communities has made the transfer of useful generic co ...

Keywords: cluster analysis, clustering applications, exploratory data analysis, incremental clustering, similarity indices, unsupervised learning

12 Array regrouping and structure splitting using whole-program reference affinity

Yutao Zhong, Maksim Orlovich, Xipeng Shen, Chen Ding
June 2004 **ACM SIGPLAN Notices , Proceedings of the ACM SIGPLAN 2004 conference on Programming language design and implementation**, Volume 39 Issue 6

Full text available: pdf(202.16 KB) Additional Information: full citation, abstract, references, citings, index terms

While the memory of most machines is organized as a hierarchy, program data are laid out in a uniform address space. This paper defines a model of *reference affinity*, which measures how close a group of data are accessed together in a reference trace. It proves that the model gives a hierarchical partition of program data. At the top is the set of all data with the weakest affinity. At the bottom is each data element with the strongest affinity. Based on the theoretical model, the paper p ...

Keywords: array regrouping, program locality, program transformation, reference affinity, reuse signature, structure splitting, volume distance

13 A survey of Web metrics



Devanshu Dhyani, Wee Keong Ng, Sourav S. Bhowmick
December 2002 **ACM Computing Surveys (CSUR)**, Volume 34 Issue 4

Full text available: pdf(289.28 KB) Additional Information: full citation, abstract, references, index terms

The unabated growth and increasing significance of the World Wide Web has resulted in a flurry of research activity to improve its capacity for serving information more effectively. But at the heart of these efforts lie implicit assumptions about "quality" and "usefulness" of Web resources and services. This observation points towards measurements and models that quantify various attributes of web sites. The science of measuring all aspects of information, especially its storage and retrieval or ...

Keywords: Information theoretic, PageRank, Web graph, Web metrics, Web page similarity, quality metrics

14 Statistics and data mining techniques for lifetime value modeling



D. R. Mani, James Drew, Andrew Betz, Piew Datta
August 1999 **Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available: pdf(1.16 MB) Additional Information: full citation, references, citings, index terms

Keywords: lifetime value, neural networks, proportional hazards regression, survival analysis, tenure prediction

15 Image Categorization by Learning and Reasoning with Regions



Yixin Chen, James Z. Wang
August 2004 **The Journal of Machine Learning Research**, Volume 5

Full text available: pdf(1.31 MB) Additional Information: full citation, abstract

Designing computer programs to automatically categorize images using low-level features is a challenging research topic in computer vision. In this paper, we present a new learning technique, which extends Multiple-Instance Learning (MIL), and its application to the problem of region-based image categorization. Images are viewed as bags, each of which contains a number of instances corresponding to regions obtained from image segmentation. The standard MIL problem assumes that a bag is labeled p ...

16 Modeling one- and two-layer variable bit rate video



Kavitha Chandra, Amy R. Reibman
June 1999 **IEEE/ACM Transactions on Networking (TON)**, Volume 7 Issue 3

Full text available: pdf(265.12 KB) Additional Information: full citation, references, citings, index terms

Keywords: MPEG2, VBR video, multiplexing, traffic model, two-layer

17 A unified framework for model-based clustering
Shi Zhong, Joydeep Ghosh
December 2003 **The Journal of Machine Learning Research**, Volume 4

Full text available: pdf(851.48 KB) Additional Information: full citation, abstract, index terms

Model-based clustering techniques have been widely used and have shown promising results in many applications involving complex data. This paper presents a unified framework for probabilistic model-based clustering based on a bipartite graph view of data and models that highlights the commonalities and differences among existing model-based clustering algorithms. In this view, clusters are represented as probabilistic models in a model space that is conceptually separate from the data space. For ...

18 Evolving data mining into solutions for insights: Scaling mining algorithms to large databases
Paul Bradley, Johannes Gehrke, Raghu Ramakrishnan, Ramakrishnan Srikant
August 2002 **Communications of the ACM**, Volume 45 Issue 8

Full text available: pdf(116.66 KB) html(28.54 KB) Additional Information: full citation, abstract, references, index terms

Which insights about data structure make it possible to analyze the very large databases collected by Internet, business, scientific, and government applications?

19 Workload models of VBR video traffic and their use in resource allocation policies
Pietro Manzoni, Paolo Cremonesi, Giuseppe Serazzi
June 1999 **IEEE/ACM Transactions on Networking (TON)**, Volume 7 Issue 3

Full text available: pdf(390.58 KB) Additional Information: full citation, references, citings, index terms

Keywords: burstiness, communication systems performance, delay-sensitive traffic, multimedia communication, networks

20 Technical session 6: learning in multi-modal data: Optimal multimodal fusion for multimedia data analysis
Yi Wu, Edward Y. Chang, Kevin Chen-Chuan Chang, John R. Smith
October 2004 **Proceedings of the 12th annual ACM international conference on Multimedia**

Full text available: pdf(350.22 KB) Additional Information: full citation, abstract, references, index terms

Considerable research has been devoted to utilizing multimodal features for better understanding multimedia data. However, two core research issues have not yet been adequately addressed. First, given a set of features extracted from multiple media sources (e.g., extracted from the visual, audio, and caption track of videos), how do we determine the best modalities? Second, once a set of modalities has been identified, how do we best fuse them to map to semantics? In this paper, we propose a ...

Keywords: curse of dimensionality, independent analysis, modality independence,

multimodal fusion, super-kernel fusion

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2005 ACM, Inc.
Terms of Usage Privacy Policy Code of Ethics Contact Us

Useful downloads: Adobe Acrobat QuickTime Windows Media Player Real Player

ACM PORTAL
US Patent & Trademark Office
Subscribe (Full Service) Register (Limited Service, Free) Login
Search: The ACM Digital Library The Guide
"k-means" and predict and margin
SEARCH
Feedback Report a problem Satisfaction survey

Terms used **k means** and **predict** and **margin** Found **3,433** of **150,138**

Sort results by: relevance
Display results: expanded form
Save results to a Binder
Search Tips
Open results in a new window
Try an Advanced Search
Try this search in The ACM Guide

Results 21 - 40 of 200 Result page: previous 1 **2** 3 4 5 6 7 8 9 10 next
Best 200 shown Relevance scale

21 Technical poster session 1: multimedia analysis, processing, and retrieval: A semi-naïve Bayesian method incorporating clustering with pair-wise constraints for auto image annotation

Wanjun Jin, Rui Shi, Tat-Seng Chua
October 2004 **Proceedings of the 12th annual ACM international conference on Multimedia**
Full text available: pdf(258.93 KB) Additional Information: full citation, abstract, references, index terms

We propose a novel approach for auto image annotation. In our approach, we first perform the segmentation of images into regions, followed by clustering of regions, before learning the relationship between concepts and region clusters using the set of training images with pre-assigned concepts. The main focus of this paper is two-fold. First, in the learning stage, we perform clustering of regions into region clusters by incorporating pair-wise constraints which are derived by considering the ...

Keywords: image annotation, pair-wise constraint, semi-naïve Bayes, semi-supervised clustering

22 Clustering algorithms: FREM: fast and robust EM clustering for large data sets

Carlos Ordonez, Edward Omiecinski
November 2002 **Proceedings of the eleventh international conference on Information and knowledge management**
Full text available: pdf(200.82 KB) Additional Information: full citation, abstract, references, citings, index terms

Clustering is a fundamental Data Mining technique. This article presents an improved EM algorithm to cluster large data sets having high dimensionality, noise and zero variance problems. The algorithm incorporates improvements to increase the quality of solutions and speed. In general the algorithm can find a good clustering solution in 3 scans over the data set. Alternatively, it can be run until it converges. The algorithm has a few parameters that are easy to set and have defaults for most ca ...

Keywords: EM, clustering, data mining

23 Text Categorization: Unsupervised document classification using sequential information maximization

Noam Slonim, Nir Friedman, Naftali Tishby
August 2002 **Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval**
Full text available: pdf(236.71 KB) Additional Information: full citation, abstract, references, index terms

We present a novel sequential clustering algorithm which is motivated by the *Information Bottleneck (IB)* method. In contrast to the agglomerative *IB* algorithm, the new sequential (*sIB*) approach is guaranteed to converge to a local maximum of the information with time and space complexity typically linear in the data size. information, as required by the original IB principle. Moreover, the time and space complexity are significantly improved. We apply this algorithm to unsup ...

24 Automating exploratory data analysis for efficient data mining
Jonathan D. Becher, Pavel Berkhin, Edmund Freeman
August 2000 **Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining**
Full text available: pdf(53.38 KB) Additional Information: full citation, references, citings, index terms

Keywords: attribute selection, automation, encoding, transformation

25 Clustering moving objects for spatio-temporal selectivity estimation
Qing Zhang, Xuemin Lin
January 2004 **Proceedings of the fifteenth conference on Australasian database - Volume 27**
Full text available: pdf(257.74 KB) Additional Information: full citation, abstract, references

Many spatio-temporal applications involve managing and querying moving objects. In such an environment, predictive spatio-temporal queries become an important query class to be processed to capture the nature of moving objects. In this paper, we investigated the problem of selectivity estimation for predictive spatio-temporal queries. We propose a novel histogram technique based on a clustering paradigm. To avoid expensive computation costs, we developed linear time heuristics to construct such ...

Keywords: histograms, predicative queries, spatio-temporal Databases

26 Trajectory clustering with mixtures of regression models
Scott Gaffney, Padhraic Smyth
August 1999 **Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining**
Full text available: pdf(1.31 MB) Additional Information: full citation, references, citings, index terms

27 Industry/government track posters: Interactive training of advanced classifiers for mining remote sensing image archives
Selim Aksoy, Krzysztof Koperski, Carsten Tusk, Giovanni Marchisio
August 2004 **Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining**
Full text available: pdf(4.24 MB) Additional Information: full citation, abstract, references, index terms

Advances in satellite technology and availability of downloaded images constantly increase the sizes of remote sensing image archives. Automatic content extraction, classification and content-based retrieval have become highly desired goals for the development of intelligent

remote sensing databases. The common approach for mining these databases uses rules created by analysts. However, incorporating GIS information and human expert knowledge with digital image processing improves remote sensing ...

Keywords: data fusion, decision tree classifiers, land cover analysis, missing data, remote sensing

28 IR-4 (information retrieval): machine learning in information retrieval: Regularizing translation models for better automatic image annotation

Feng Kang, Rong Jin, Joyce Y. Chai

November 2004 **Proceedings of the Thirteenth ACM conference on Information and knowledge management**

Full text available: pdf(250.23 KB) Additional Information: full citation, abstract, references, index terms

The goal of automatic image annotation is to automatically generate annotations for images to describe their content. In the past, statistical machine translation models have been successfully applied to automatic image annotation task [8]. It views the process of annotating images as a process of translating the content from a 'visual language' to textual words. One problem with the existing translation models is that common words are usually associated with too many different image regions. ...

Keywords: automatic image annotation, normalized translation model, regularized translation model, translation model

29 Probabilistic hierarchical clustering for biological data

Eran Segal, Daphne Koller

April 2002 **Proceedings of the sixth annual international conference on Computational biology**

Full text available: pdf(2.06 MB) Additional Information: full citation, abstract, citings, index terms

Biological data, such as gene expression profiles or protein sequences, is often organized in a hierarchy of classes, where the instances assigned to "nearby" classes in the tree are similar. Most approaches for constructing a hierarchy use simple local operations, that are very sensitive to noise or variation in the data. In this paper, we describe probabilistic abstraction hierarchies (PAH) [11], a general probabilistic framework for clustering data into a hierarchy, and show how it can be app ...

30 Active learning using pre-clustering

Hieu T. Nguyen, Arnold Smeulders

July 2004 **Twenty-first international conference on Machine learning**

Full text available: pdf(166.97 KB) Additional Information: full citation, abstract, references, citings

The paper is concerned with two-class active learning. While the common approach for collecting data in active learning is to select samples close to the classification boundary, better performance can be achieved by taking into account the prior data distribution. The main contribution of the paper is a formal framework that incorporates clustering into active learning. The algorithm first constructs a classifier on the set of the cluster representatives, and then propagates the classification ...

31 Temporal sequence learning and data reduction for anomaly detection

Terran Lane, Carla E. Brodley

August 1999 **ACM Transactions on Information and System Security (TISSEC)**, Volume 2 Issue 3

Full text available: pdf(628.31 KB) Additional Information: full citation, abstract, references, citings, index

terms

The anomaly-detection problem can be formulated as one of learning to characterize the behaviors of an individual, system, or network in terms of temporal sequences of discrete data. We present an approach on the basis of instance-based learning (IBL) techniques. To cast the anomaly-detection task in an IBL framework, we employ an approach that transforms temporal sequences of discrete, unordered observations into a metric space via a similarity measure that encodes intra-attribute depende ...

Keywords: anomaly detection, clustering, data reduction, empirical evaluation, instance based learning, machine learning, user profiling

32 SONIA: a service for organizing networked information autonomously
Mehran Sahami, Salim Yusufali, Michelle Q. W. Baldonaldo
May 1998 **Proceedings of the third ACM conference on Digital libraries**
Full text available: pdf(1.29 MB) Additional Information: full citation, references, citings, index terms

33 Reducing multiclass to binary: a unifying approach for margin classifiers
Erin L. Allwein, Robert E. Schapire, Yoram Singer
September 2001 **The Journal of Machine Learning Research**, Volume 1
Full text available: pdf(310.85 KB) Additional Information: full citation, abstract

We present a unifying framework for studying the solution of multiclass categorization problems by reducing them to multiple binary problems that are then solved using a margin-based binary learning algorithm. The proposed framework unifies some of the most popular approaches in which each class is compared against all others, or in which all pairs of classes are compared to each other, or in which output codes with error-correcting properties are used. We propose a general method for combining ...

34 Cluster-based find and replace
Robert C. Miller, Alisa M. Marshall
April 2004 **Proceedings of the 2004 conference on Human factors in computing systems**
Full text available: pdf(190.25 KB) Additional Information: full citation, abstract, references, index terms

In current text editors, the find & replace command offers only two options: replace one match at a time prompting for confirmation, or replace all matches at once without any confirmation. Both approaches are prone to errors. This paper explores a third way: *cluster-based find & replace*, in which the matches are clustered by similarity and whole clusters can be replaced at once. We hypothesized that cluster-based find & replace would make find & replace tasks both faster and more accurat ...

Keywords: clustering, error prevention, find & replace, text editing

35 Multimedia information retrieval: Automatic image annotation and retrieval using cross-media relevance models
J. Jeon, V. Lavrenko, R. Manmatha
July 2003 **Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval**
Full text available: pdf(539.83 KB) Additional Information: full citation, abstract, references, citings, index terms

Libraries have traditionally used manual image annotation for indexing and then later

retrieving their image collections. However, manual image annotation is an expensive and labor intensive procedure and hence there has been great interest in coming up with automatic ways to retrieve images based on content. Here, we propose an automatic approach to annotating and retrieving images based on a training set of images. We assume that regions in an image can be described using a small vocabulary of ...

Keywords: image annotation, image retrieval, relevance models

36 Simulation-based planning for multi-agent environments

Jin Joo Lee, Paul A. Fishwick

December 1997 **Proceedings of the 29th conference on Winter simulation**

Full text available: pdf(997.79 KB) Additional Information: full citation, references, index terms

37 A study of scalar compilation techniques for pipelined supercomputers

Shlomo Weiss, James E. Smith

September 1990 **ACM Transactions on Mathematical Software (TOMS)**, Volume 16 Issue 3

Full text available: pdf(1.47 MB) Additional Information: full citation, abstract, references, index terms, review

This paper studies two compilation techniques for enhancing scalar performance in high-speed scientific processors: software pipelining and loop unrolling. We study the impact of the architecture (size of the register file) and of the hardware (size of instruction buffer) on the efficiency of loop unrolling. We also develop a methodology for classifying software pipelining techniques. For loop unrolling, a straightforward scheduling algorithm is shown to produce near-optimal results when no ...

38 Bringing order to the Web: automatically categorizing search results

Hao Chen, Susan Dumais

April 2000 **Proceedings of the SIGCHI conference on Human factors in computing systems**

Full text available: pdf(1.00 MB) Additional Information: full citation, abstract, references, citings, index terms

We developed a user interface that organizes Web search results into hierarchical categories. Text classification algorithms were used to automatically classify arbitrary search results into an existing category structure on-the-fly. A user study compared our new category interface with the typical ranked list interface of search results. The study showed that the category interface is superior both in objective and subjective measures. Subjects liked the category interface much better than t ...

Keywords: World Wide Web, classification, search, support vector machine, text categorization, text categrization, user interface, user study

39 Large margin classification using the perceptron algorithm

Yoav Freund, Robert E. Schapire

July 1998 **Proceedings of the eleventh annual conference on Computational learning theory**

Full text available: pdf(1.11 MB) Additional Information: full citation, references, citings, index terms

40 Large margin hierarchical classification

Ofer Dekel, Joseph Keshet, Yoram Singer
July 2004 **Twenty-first international conference on Machine learning**

Full text available: pdf(126.37 KB) Additional Information: full citation, abstract, references

We present an algorithmic framework for supervised classification learning where the set of labels is organized in a predefined hierarchical structure. This structure is encoded by a rooted tree which induces a metric over the label set. Our approach combines ideas from large margin kernel methods and Bayesian analysis. Following the large margin principle, we associate a prototype with each label in the tree and formulate the learning task as an optimization problem with varying margin constrai ...

Results 21 - 40 of 200 Result page: previous 1 **2** 3 4 5 6 7 8 9 10 next

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2005 ACM, Inc.
Terms of Usage Privacy Policy Code of Ethics Contact Us

Useful downloads: Adobe Acrobat QuickTime Windows Media Player Real Player

PORTAL
US Patent & Trademark Office

Subscribe (Full Service) Register (Limited Service, Free) Login

Search: The ACM Digital Library The Guide

"k-means" and predict and margin SEARCH

Feedback Report a problem Satisfaction survey

Terms used **k means** and **predict** and **margin** Found **3,433** of **150,138**

Sort results by relevance

Display results expanded form

Save results to a Binder

Search Tips

Open results in a new window

Try an Advanced Search

Try this search in The ACM Guide

Results 41 - 60 of 200 Result page: previous 1 2 **3** 4 5 6 7 8 9 10 next

Best 200 shown Relevance scale

41 Improved boosting algorithms using confidence-rated predictions

Robert E. Schapire, Yoram Singer

July 1998 **Proceedings of the eleventh annual conference on Computational learning theory**

Full text available: pdf(1.59 MB) Additional Information: full citation, references, citings, index terms

42 Industrial/government track: Frequent-subsequence-based prediction of outer membrane proteins

Rong She, Fei Chen, Ke Wang, Martin Ester, Jennifer L. Gardy, Fiona S. L. Brinkman

August 2003 **Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available: pdf(166.07 KB) Additional Information: full citation, abstract, references, index terms

A number of medically important disease-causing bacteria (collectively called Gram-negative bacteria) are noted for the extra "outer" membrane that surrounds their cell. Proteins resident in this membrane (outer membrane proteins, or OMPs) are of primary research interest for antibiotic and vaccine drug design as they are on the surface of the bacteria and so are the most accessible targets to develop new drugs against. With the development of genome sequencing technology and bioinformatics, bio ...

Keywords: association rule, classification, outer membrane protein, subcellular localization, support vector machine

43 Special issue on kernel methods: A new approximate maximal margin classification algorithm

Claudio Gentile

March 2002 **The Journal of Machine Learning Research**, Volume 2

Full text available: pdf(361.70 KB) Additional Information: full citation, abstract, citings

A new incremental learning algorithm is described which approximates the maximal margin hyperplane w.r.t. norm $p \geq 2$ for a set of linearly separable data. Our algorithm, called ALMA_p (Approximate Large Margin algorithm w.r.t. norm p), takes $O((p-1) / (\alpha^2 \gamma^2))$ corrections to separate the data with p-norm margin larger than $(1-\alpha)\gamma$, where g is the (normalized) p-norm margin of the data. ALMA ...

44 Static correlated branch prediction

Cliff Young, Michael D. Smith

September 1999 **ACM Transactions on Programming Languages and Systems (TOPLAS)**, Volume 21 Issue 5

Full text available: pdf(508.49 KB) Additional Information: full citation, abstract, references, citings, index terms

Recent work in history-based branch prediction uses novel hardware structures to capture branch correlation and increase branch prediction accuracy. Branch correlation occurs when the outcome of a conditional branch can be accurately predicted by observing the outcomes of previously executed branches in the dynamic instruction stream. In this article, we show how to instrument a program so that it is practical to collect run-time statistics that indicate where branch correl ...

Keywords: branch correlation, branch prediction, path profiling, profile-driven optimization

45 Information retrieval session 4: general retrieval issues I: Margin-based local regression for adaptive filtering

Yiming Yang, Bryan Kisiel

November 2003 **Proceedings of the twelfth international conference on Information and knowledge management**

Full text available: pdf(2.23 MB) Additional Information: full citation, abstract, references, index terms

Adaptive information filtering is an open challenge in information retrieval. One of the tough issues is the optimization of decision thresholds over time, based on partial relevance feedback on the system-retrieved documents in chronological order. We developed a new approach, namely margin-based local regression, that automatically adjusts the thresholds based on a sliding window over the truly positive examples for which the system predicted "yes" with respect to a particular class, and a sec ...

Keywords: adaptive filtering, local regression, temporal sequences, threshold calibration

46 Boosting as a Regularized Path to a Maximum Margin Classifier

Saharon Rosset, Ji Zhu, Trevor Hastie

August 2004 **The Journal of Machine Learning Research**, Volume 5

Full text available: pdf(553.71 KB) Additional Information: full citation, abstract

In this paper we study boosting methods from a new perspective. We build on recent work by Efron et al. to show that boosting approximately (and in some cases exactly) minimizes its loss criterion with an l_1 constraint on the coefficient vector. This helps understand the success of boosting with early stopping as regularized fitting of the loss criterion. For the two most commonly used criteria (exponential and binomial log-likelihood), we further show that as the constraint is ...

47 Statistical Analysis of Some Multi-Category Large Margin Classification Methods

Tong Zhang

December 2004 **The Journal of Machine Learning Research**, Volume 5

Full text available: pdf(244.80 KB) Additional Information: full citation, abstract

The purpose of this paper is to investigate statistical properties of risk minimization based multi-category classification methods. These methods can be considered as natural extensions of binary large margin classification. We establish conditions that guarantee the consistency of classifiers obtained in the risk minimization framework with respect to the classification error. Examples are provided for four specific forms of the general

formulation, which extend a number of known methods. Usin ...

48 Ultraconservative online algorithms for multiclass problems
Koby Crammer, Yoram Singer
March 2003 **The Journal of Machine Learning Research**, Volume 3
Full text available: pdf(255.98 KB) Additional Information: full citation, abstract, index terms

In this paper we study a paradigm to generalize online classification algorithms for binary classification problems to multiclass problems. The particular hypotheses we investigate maintain one prototype vector per class. Given an input instance, a multiclass hypothesis computes a similarity-score between each prototype and the input instance and sets the predicted label to be the index of the prototype achieving the highest similarity. To design and analyze the learning algorithms in this paper ...

49 Research track papers: Incorporating prior knowledge with weighted margin support vector machines
Xiaoyun Wu, Rohini Srihari
August 2004 **Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining**
Full text available: pdf(184.84 KB) Additional Information: full citation, abstract, references, index terms

Like many purely data-driven machine learning methods, Support Vector Machine (SVM) classifiers are learned exclusively from the evidence presented in the training dataset; thus a larger training dataset is required for better performance. In some applications, there might be human knowledge available that, in principle, could compensate for the lack of data. In this paper, we propose a simple generalization of SVM: Weighted Margin SVM (WMSVMs) that permits the incorporation of prior knowledge. ...

Keywords: incorporating prior knowledge, support vector machines, text categorization

50 Classification: SVM binary classifier ensembles for image classification
King-Shy Goh, Edward Chang, Kwang-Ting Cheng
October 2001 **Proceedings of the tenth international conference on Information and knowledge management**
Full text available: pdf(1.80 MB) Additional Information: full citation, abstract, references, citings, index terms

We study how the SVM-based binary classifiers can be effectively combined to tackle the multi-class image classification problem. We study several ensemble schemes, including OPC (one per class), PWC (pairwise coupling), and ECOC (error-correction output coding), that aim to achieve good error correction capability through redundancy. To enhance these ensemble schemes' accuracy, we propose methods that on the one hand boost the margins (i.e., confidence) of the SVM-based binary classifiers, and, ...

51 Information access and retrieval (IAR): A comparison of several predictive algorithms for collaborative filtering on multi-valued ratings
Maritza L. Calderón-Benavides, Cristina N. González-Caro, José de J. Pérez-Alcázar, Juan C. García-Díaz, Joaquin Delgado
March 2004 **Proceedings of the 2004 ACM symposium on Applied computing**
Full text available: pdf(193.01 KB) Additional Information: full citation, abstract, references, index terms

The basic objective of a predictive algorithm for collaborative filtering (CF) is to suggest items to a particular user based on his/her preferences and other users with similar interests. Many algorithms have been proposed for CF, and some works comparing sub-sets of them can be found in the literature; however, more comprehensive comparisons are not

available. In this work, a meaningful sample of CF algorithms widely reported in the literature were chosen for analysis; they represent different ...

Keywords: Dependency Networks, Support Vector Machines, collaborative filtering, memory-based models, online learning

52 Pac-bayesian generalisation error bounds for gaussian process classification
Matthias Seeger
March 2003 **The Journal of Machine Learning Research**, Volume 3

Full text available: pdf(487.11 KB) Additional Information: full citation, abstract, references, index terms

Approximate Bayesian Gaussian process (GP) classification techniques are powerful non-parametric learning methods, similar in appearance and performance to support vector machines. Based on simple probabilistic models, they render interpretable results and can be embedded in Bayesian frameworks for model selection, feature selection, etc. In this paper, by applying the PAC-Bayesian theorem of McAllester (1999a), we prove distribution-free generalisation error bounds for a wide range of approxima ...

Keywords: Bayesian learning, Gaussian processes, Gibbs classifier, Kernel machines, PAC-Bayesian framework, convex duality, generalisation error bounds, sparse approximations

53 Sparse bayesian learning and the relevance vector machine
Michael E. Tipping
September 2001 **The Journal of Machine Learning Research**, Volume 1

Full text available: pdf(999.88 KB) Additional Information: full citation, abstract

This paper introduces a general Bayesian framework for obtaining sparse solutions to regression and classification tasks utilising models linear in the parameters. Although this framework is fully general, we illustrate our approach with a particular specialisation that we denote the 'relevance vector machine' (RVM), a model of identical functional form to the popular and state-of-the-art 'support vector machine' (SVM). We demonstrate that by exploiting a probabilistic Bayesian learning framewor ...

54 Class prediction and discovery using gene expression data
Donna K. Slonim, Pablo Tamayo, Jill P. Mesirov, Todd R. Golub, Eric S. Lander
April 2000 **Proceedings of the fourth annual international conference on Computational molecular biology**

Full text available: pdf(858.00 KB) Additional Information: full citation, abstract, references, citings

Classification of patient samples is a crucial aspect of cancer diagnosis and treatment. We present a method for classifying samples by computational analysis of gene expression data. We consider the classification problem in two parts: *class discovery* and *class prediction*. Class discovery refers to the process of dividing samples into reproducible classes that have similar behavior or properties, while class prediction places new samples into already known classes. We describe ...

55 A graphical model for protein secondary structure prediction
Wei Chu, Zoubin Ghahramani, David L. Wild
July 2004 **Twenty-first international conference on Machine learning**

Full text available: pdf(366.19 KB) Additional Information: full citation, abstract, references

In this paper, we present a graphical model for protein secondary structure prediction. This model extends segmental semi-Markov models (SSMM) to exploit multiple sequence alignment profiles which contain information from evolutionarily related sequences. A novel

parameterized model is proposed as the likelihood function for the SSMM to capture the segmental conformation. By incorporating the information from long range interactions in ß-sheets, this model is capable of carrying out infere ...

56 A training algorithm for optimal margin classifiers
Bernhard E. Boser, Isabelle M. Guyon, Vladimir N. Vapnik
July 1992 **Proceedings of the fifth annual workshop on Computational learning theory**

Full text available: pdf(1.00 MB) Additional Information: full citation, abstract, references, citings, index terms

A training algorithm that maximizes the margin between the training patterns and the decision boundary is presented. The technique is applicable to a wide variety of the classification functions, including Perceptrons, polynomials, and Radial Basis Functions. The effective number of parameters is adjusted automatically to match the complexity of the problem. The solution is expressed as a linear combination of supporting patterns. These are the subset of training patterns that are closest t ...

57 Predictive automatic relevance determination by expectation propagation
Yuan (Alan) Qi, Thomas P. Minka, Rosalind W. Picard, Zoubin Ghahramani
July 2004 **Twenty-first international conference on Machine learning**

Full text available: pdf(314.64 KB) Additional Information: full citation, abstract, references

In many real-world classification problems the input contains a large number of potentially irrelevant features. This paper proposes a new Bayesian framework for determining the relevance of input features. This approach extends one of the most successful Bayesian methods for feature selection and sparse learning, known as Automatic Relevance Determination (ARD). ARD finds the relevance of features by optimizing the model marginal likelihood, also known as the evidence. We show that this can lea ...

58 Session 6: Battery lifetime prediction for energy-aware computing
Daler Rakhmatov, Sarma Vrudhula, Deborah A. Wallach
August 2002 **Proceedings of the 2002 international symposium on Low power electronics and design**

Full text available: pdf(108.94 KB) Additional Information: full citation, abstract, references, citings, index terms

Predicting the time of full discharge of a finite-capacity energy source, such as a battery, is important for the design of portable electronic systems and applications. In this paper we present a novel analytical model of a battery that not only can be used to predict battery lifetime, but also can serve as a cost function for optimization of the energy usage in battery-powered systems. The model is physically justified, and involves only two parameters, which are easily estimated. The paper in ...

Keywords: battery, low-power design, modeling

59 Trading agents: Walverine: a Walrasian trading agent
Shih-Fen Cheng, Evan Leung, Kevin M. Lochner, Kevin O'Malley, Daniel M. Reeves, L. Julian Schvartzman, Michael P. Wellman
July 2003 **Proceedings of the second international joint conference on Autonomous agents and multiagent systems**

Full text available: pdf(151.73 KB) Additional Information: full citation, abstract, references, citings, index terms

TAC-02 was the third in a series of Trading Agent Competition events fostering research in automating trading strategies by showcasing alternate approaches in an open-invitation market game. TAC presents a challenging travel-shopping scenario where agents must

satisfy client preferences for complementary and substitutable goods by interacting through a variety of market types. Michigan's entry, Walverine, attempts to bid optimally based on a competitive analysis of the TAC travel economy. Walver ...

Keywords: competitive equilibrium, trading agents

60 Special issue on Machine learning methods for text and images: A family of additive online algorithms for category ranking

Koby Crammer, Yoram Singer

March 2003 **The Journal of Machine Learning Research**, Volume 3

Full text available: pdf(1.19 MB) Additional Information: full citation, abstract, index terms

We describe a new family of topic-ranking algorithms for multi-labeled documents. The motivation for the algorithms stem from recent advances in online learning algorithms. The algorithms are simple to implement and are also time and memory efficient. We provide a unified analysis of the family of algorithms in the mistake bound model. We then discuss experiments with the proposed family of topic-ranking algorithms on the Reuters-21578 corpus and the new corpus released by Reuters in 2000. On bo ...

Results 41 - 60 of 200 Result page: previous 1 2 **3** 4 5 6 7 8 9 10 next

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2005 ACM, Inc.
Terms of Usage Privacy Policy Code of Ethics Contact Us

Useful downloads: Adobe Acrobat QuickTime Windows Media Player Real Player

USPTO Patent Full-Text and Image Database

Searching 1976 to present...

Results of Search in 1976 to present db for:
(("k-means" AND predict) AND margin): 6 patents.
Hits 1 through 6 out of 6

Jump To

Refine Search "k-means" and predict and margin

	PAT. NO.	Title
1	6,834,109	Method and apparatus for mitigation of disturbers in communication systems
2	6,753,142	Withdrawn
3	6,647,341	Methods for classifying samples and ascertaining previously unknown classes
4	6,084,547	Enhanced position calculation
5	5,917,449	Enhanced position calculation
6	5,717,406	Enhanced position calculation